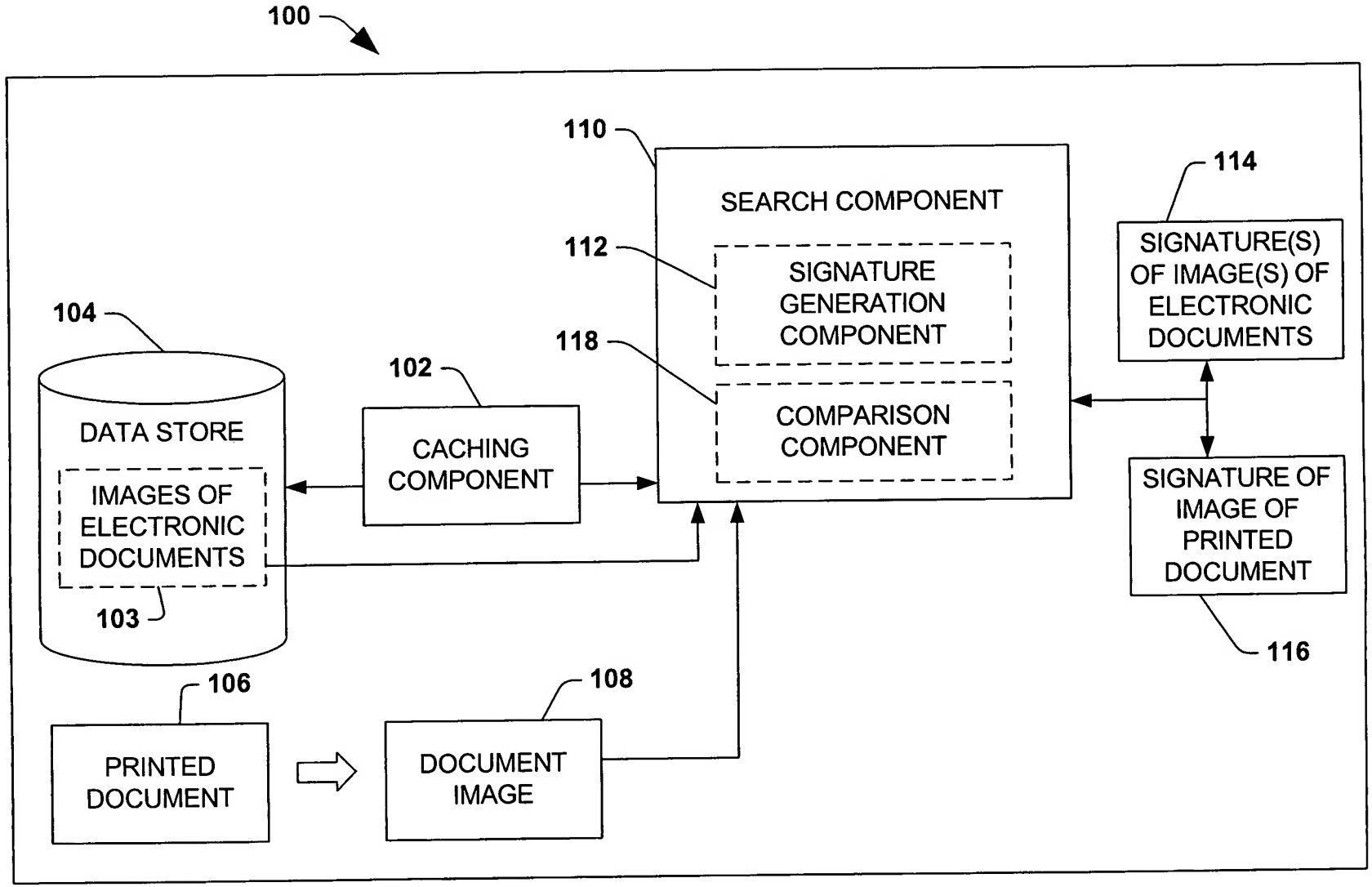
100
110
SEARCH COMPONENT
112
SIGNATURE GENERATION COMPONENT
118
COMPARISON COMPONENT
114
SIGNATURE(S) OF IMAGE(S) OF ELECTRONIC DOCUMENTS
104
DATA STORE
IMAGES OF ELECTRONIC DOCUMENTS
103
102
CACHING COMPONENT
SIGNATURE OF IMAGE OF PRINTED DOCUMENT
116
106
PRINTED DOCUMENT
108
DOCUMENT IMAGE

FIG. 1

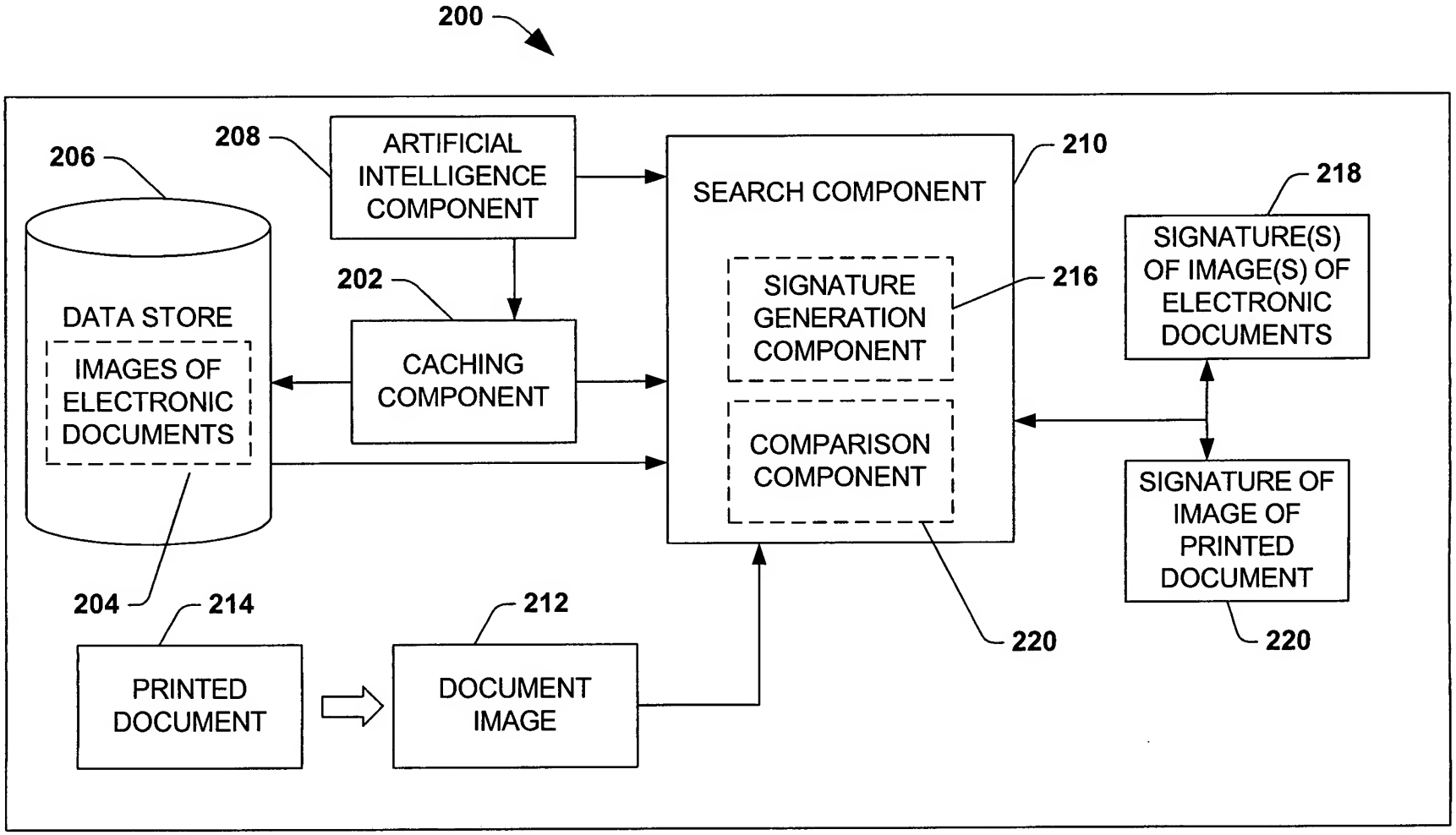
200
206
208
ARTIFICIAL INTELLIGENCE COMPONENT
210
SEARCH COMPONENT
218
SIGNATURE(S) OF IMAGE(S) OF ELECTRONIC DOCUMENTS
202
DATA STORE
IMAGES OF ELECTRONIC DOCUMENTS
CACHING COMPONENT
SIGNATURE GENERATION COMPONENT
216
COMPARISON COMPONENT
SIGNATURE OF IMAGE OF PRINTED DOCUMENT
204
214
212
220
220
PRINTED DOCUMENT
DOCUMENT IMAGE

FIG. 2

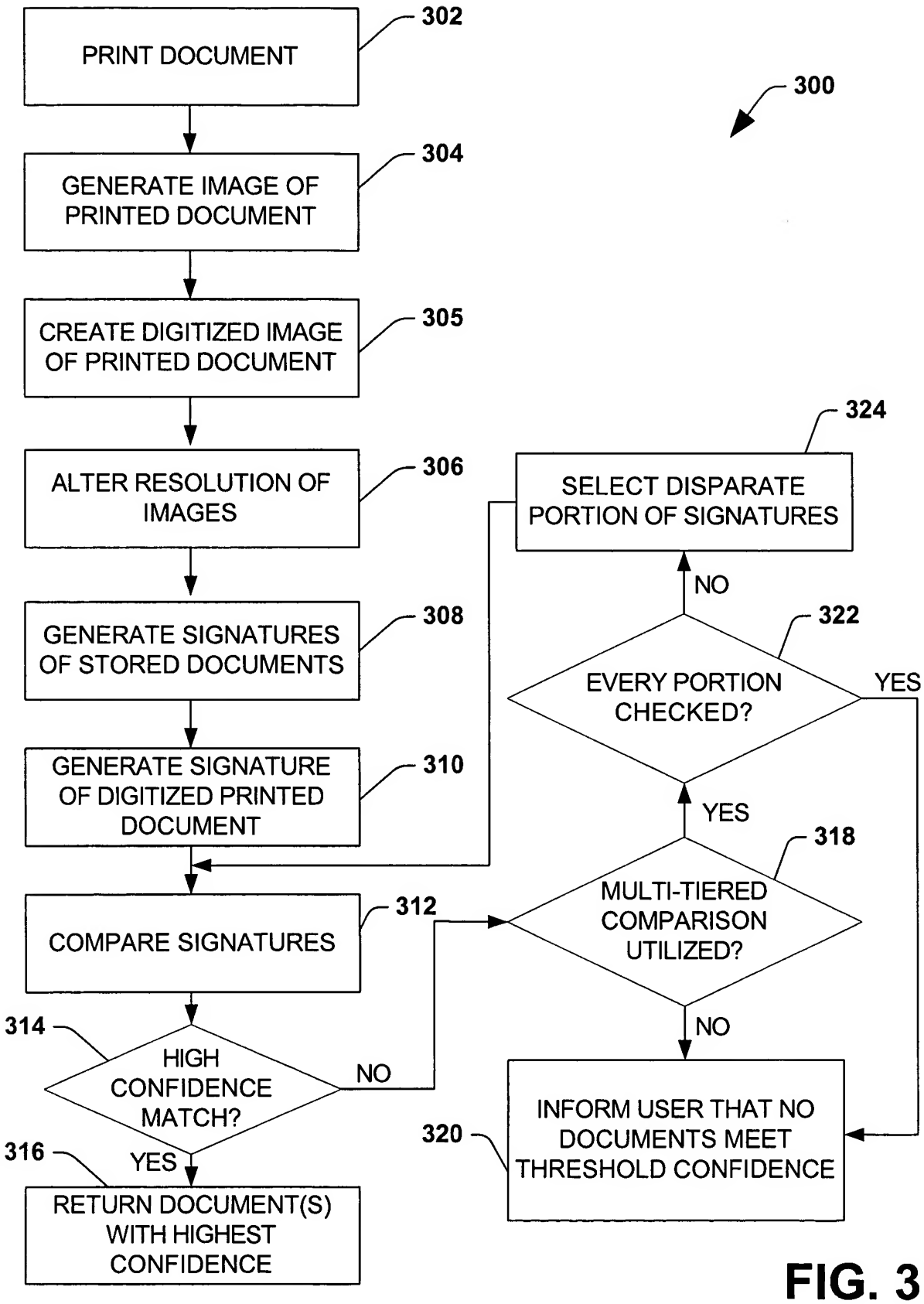
PRINT DOCUMENT
302
300
GENERATE IMAGE OF PRINTED DOCUMENT
304
CREATE DIGITIZED IMAGE OF PRINTED DOCUMENT
305
ALTER RESOLUTION OF IMAGES
306
GENERATE SIGNATURES OF STORED DOCUMENTS
308
GENERATE SIGNATURE OF DIGITIZED PRINTED DOCUMENT
310
COMPARE SIGNATURES
312
HIGH CONFIDENCE MATCH?
314
NO
YES
316
RETURN DOCUMENT(S) WITH HIGHEST CONFIDENCE
MULTI-TIERED COMPARISON UTILIZED?
318
NO
YES
EVERY PORTION CHECKED?
322
NO
YES
SELECT DISPARATE PORTION OF SIGNATURES
324
INFORM USER THAT NO DOCUMENTS MEET THRESHOLD CONFIDENCE
320

FIG. 3

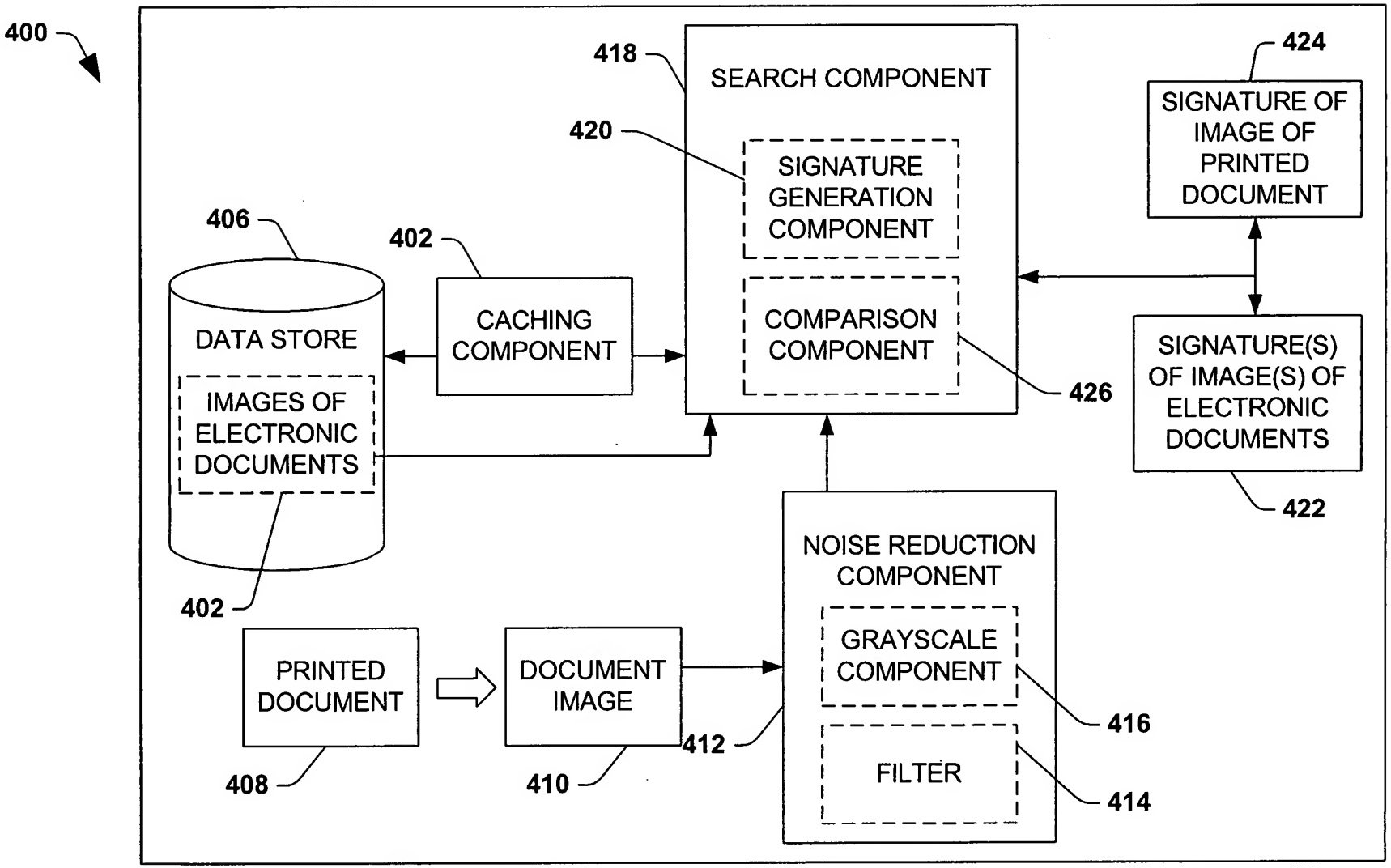
400
418
SEARCH COMPONENT
420
SIGNATURE GENERATION COMPONENT
COMPARISON COMPONENT
426
424
SIGNATURE OF IMAGE OF PRINTED DOCUMENT
406
402
DATA STORE
IMAGES OF ELECTRONIC DOCUMENTS
CACHING COMPONENT
SIGNATURE(S) OF IMAGE(S) OF ELECTRONIC DOCUMENTS
422
402
NOISE REDUCTION COMPONENT
GRAYSCALE COMPONENT
416
FILTER
414
412
PRINTED DOCUMENT
DOCUMENT IMAGE
408
410

FIG. 4

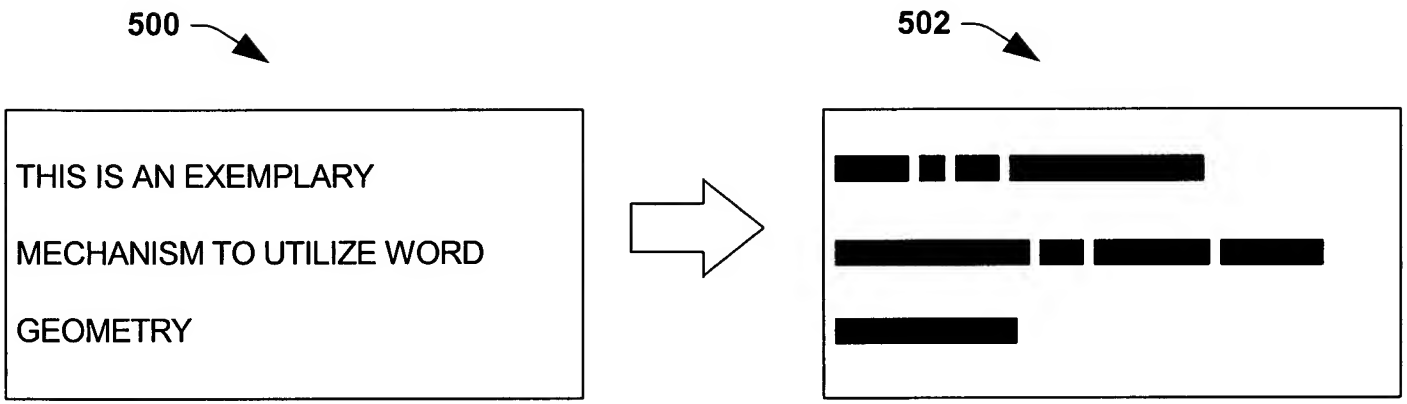
500
502
THIS IS AN EXEMPLARY
MECHANISM TO UTILIZE WORD
GEOMETRY

FIG. 5

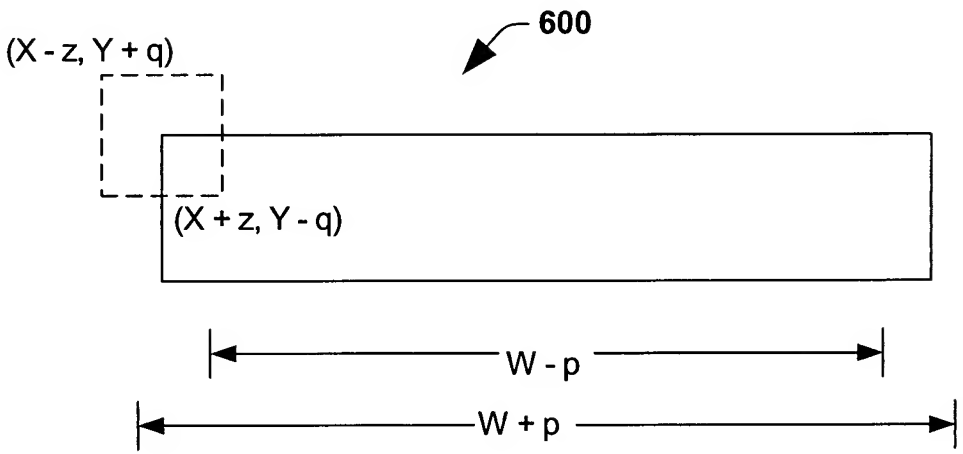
600
(X - z, Y + q)
(X + z, Y - q)
W - p
W + p

FIG. 6

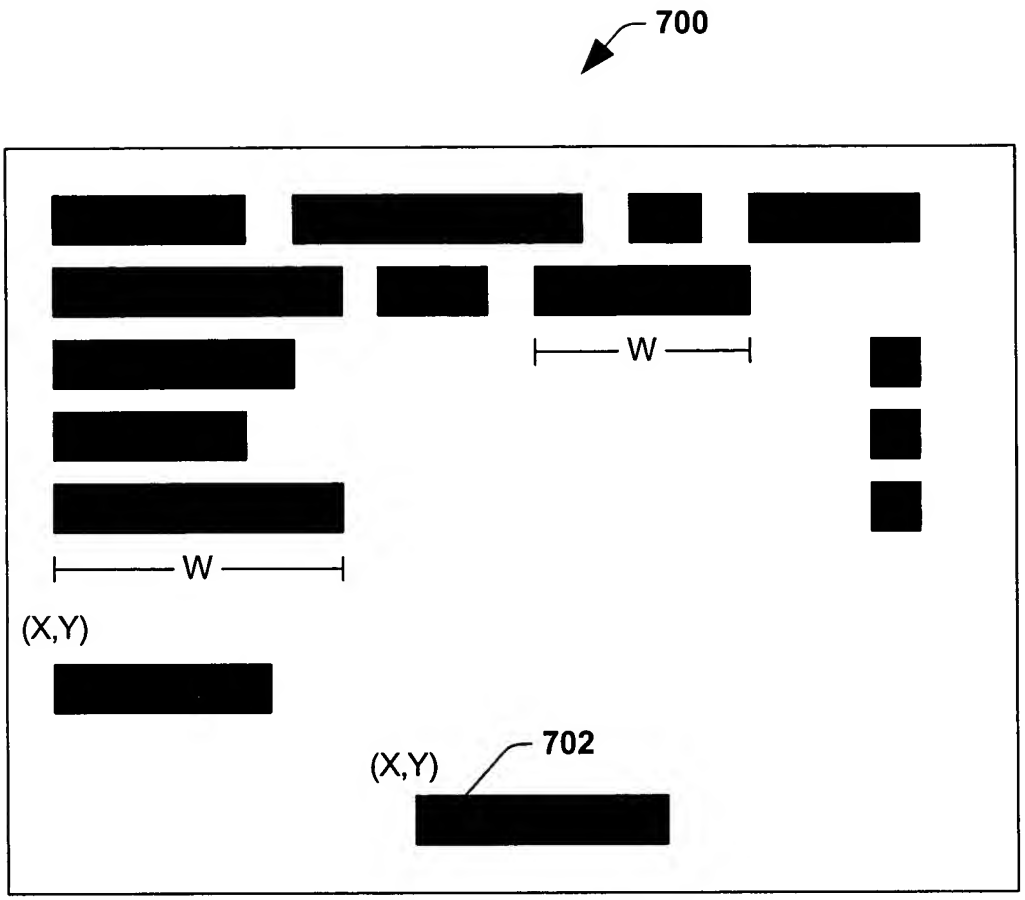

FIG. 7

BEST AVAILABLE COPY

800

802 804

H(X,Y,W)	VALUE
H(31, 21, 16)	FALSE
H(31, 21, 17)	TRUE
… … … …	… … … …
H(33, 23, 19)	TRUE
H(33, 23, 20)	FALSE
… … … …	… … … …
H(50, 16, 12)	FALSE
H(50, 16, 13)	TRUE
… … … …	… … … …
H(52, 18, 15)	TRUE
H(52, 18, 16)	FALSE

FIG. 8

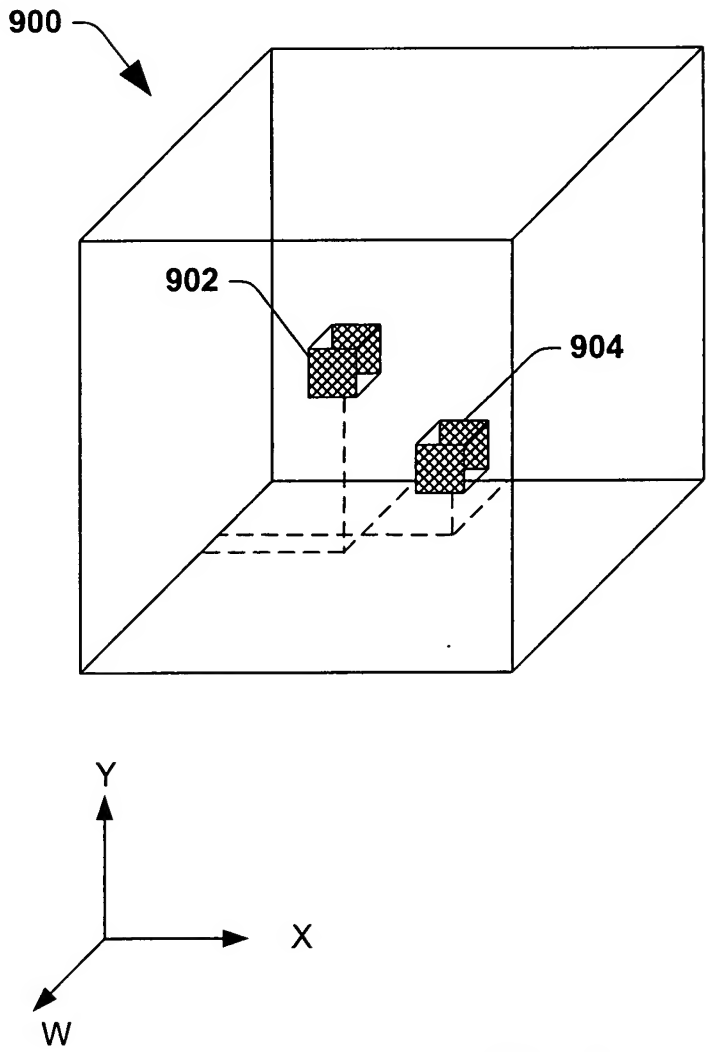

FIG. 9

BEST AVAILABLE COPY

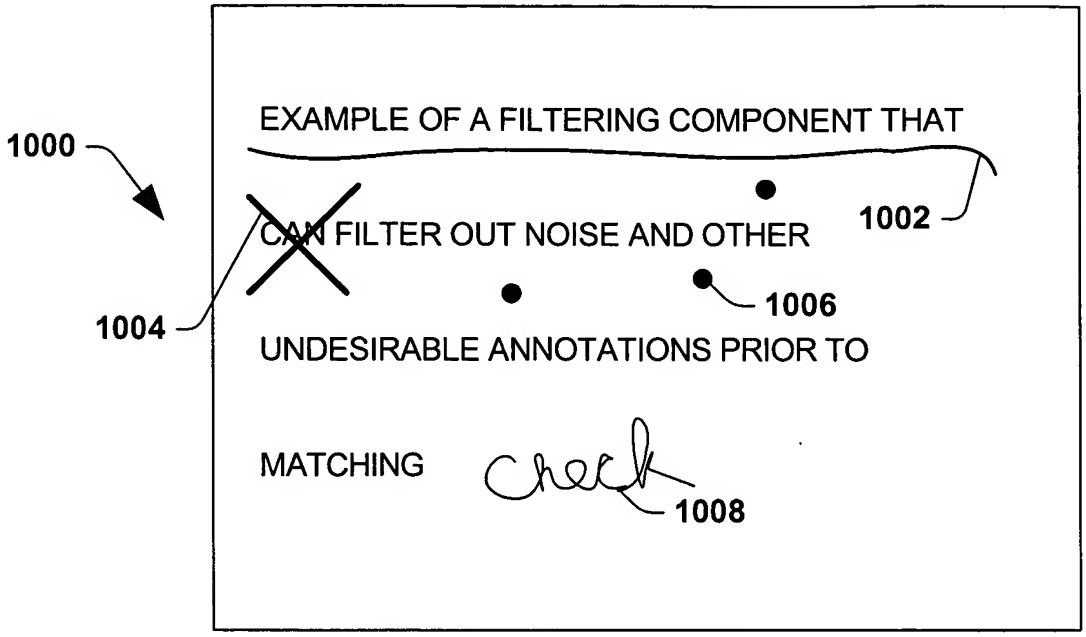

FIG. 10

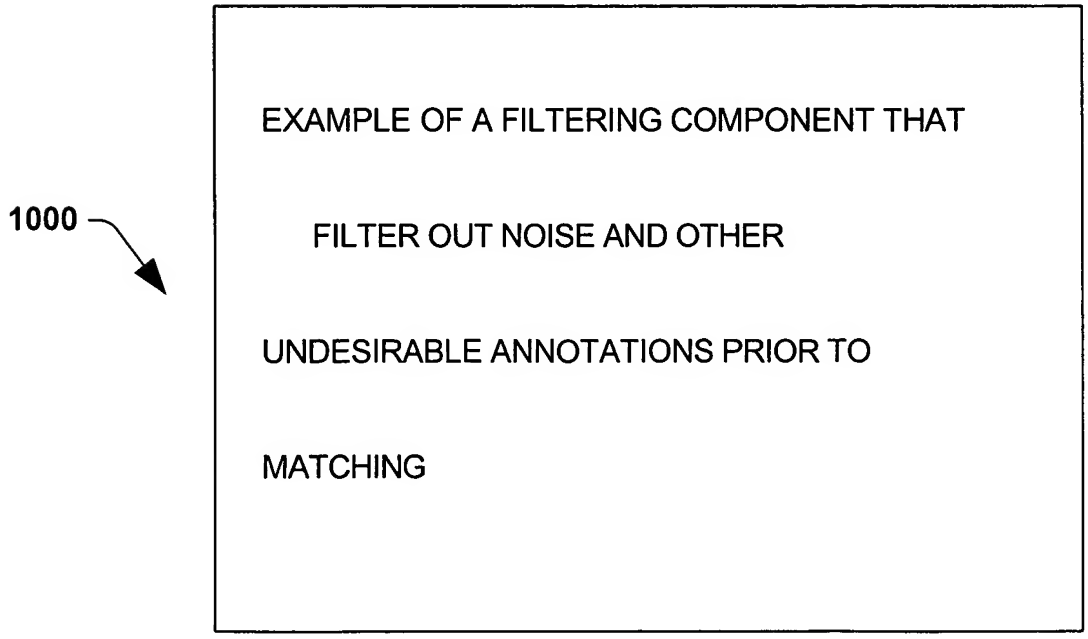

FIG. 11

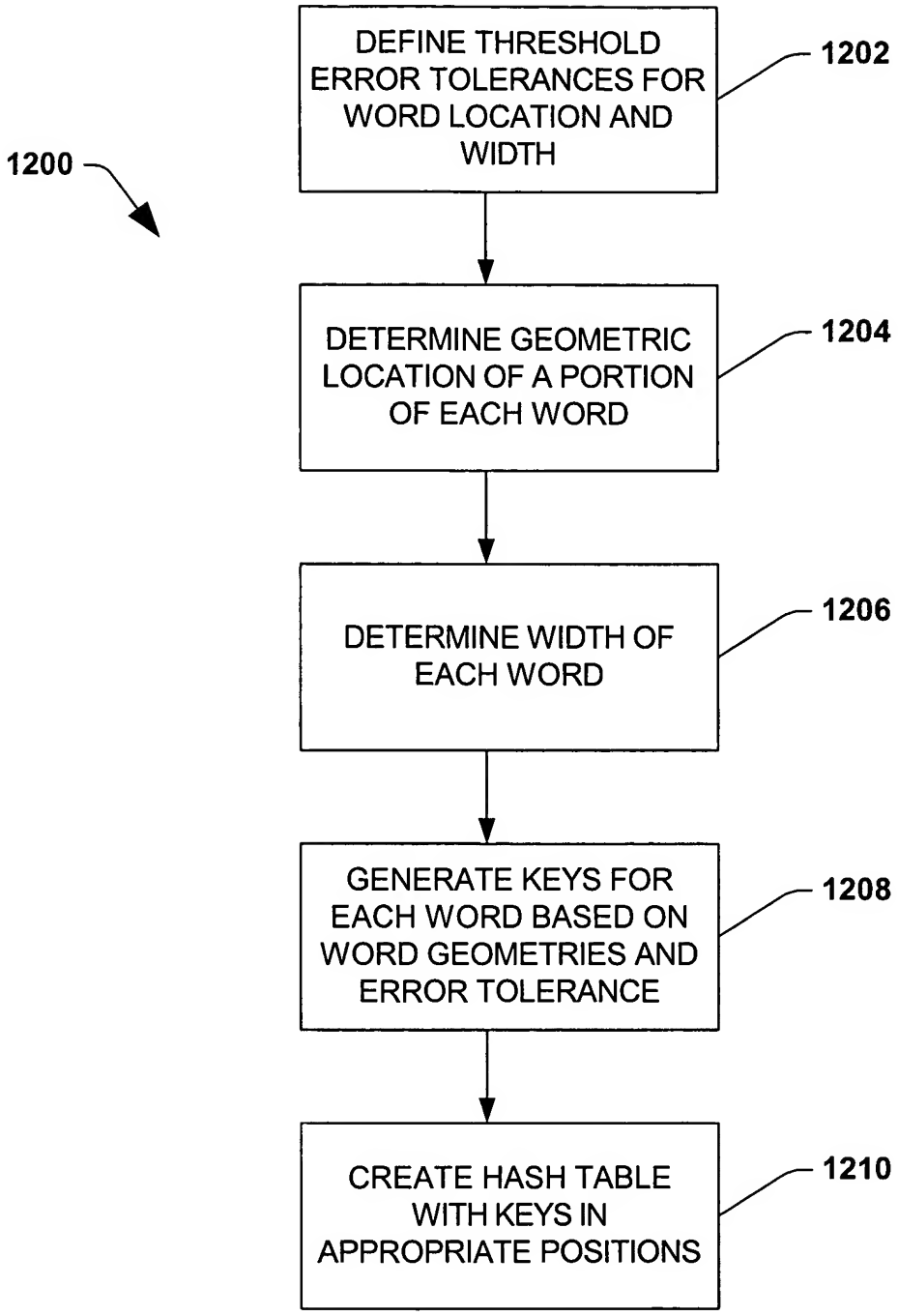
1200
DEFINE THRESHOLD ERROR TOLERANCES FOR WORD LOCATION AND WIDTH
1202
DETERMINE GEOMETRIC LOCATION OF A PORTION OF EACH WORD
1204
DETERMINE WIDTH OF EACH WORD
1206
GENERATE KEYS FOR EACH WORD BASED ON WORD GEOMETRIES AND ERROR TOLERANCE
1208
CREATE HASH TABLE WITH KEYS IN APPROPRIATE POSITIONS
1210

FIG. 12

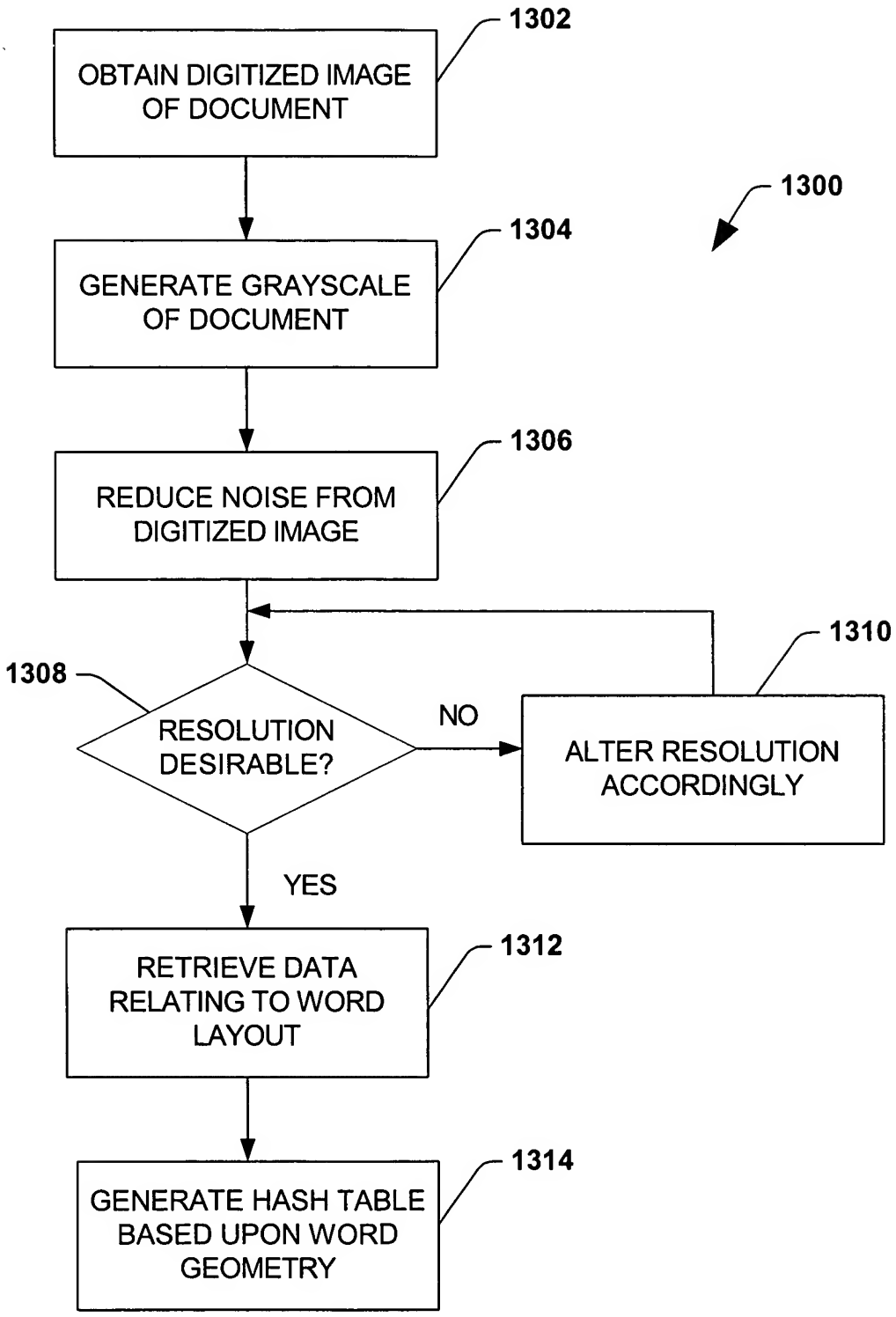
1302
OBTAIN DIGITIZED IMAGE OF DOCUMENT
1300
1304
GENERATE GRAYSCALE OF DOCUMENT
1306
REDUCE NOISE FROM DIGITIZED IMAGE
1310
1308
RESOLUTION DESIRABLE?
NO
ALTER RESOLUTION ACCORDINGLY
YES
1312
RETRIEVE DATA RELATING TO WORD LAYOUT
1314
GENERATE HASH TABLE BASED UPON WORD GEOMETRY

FIG. 13

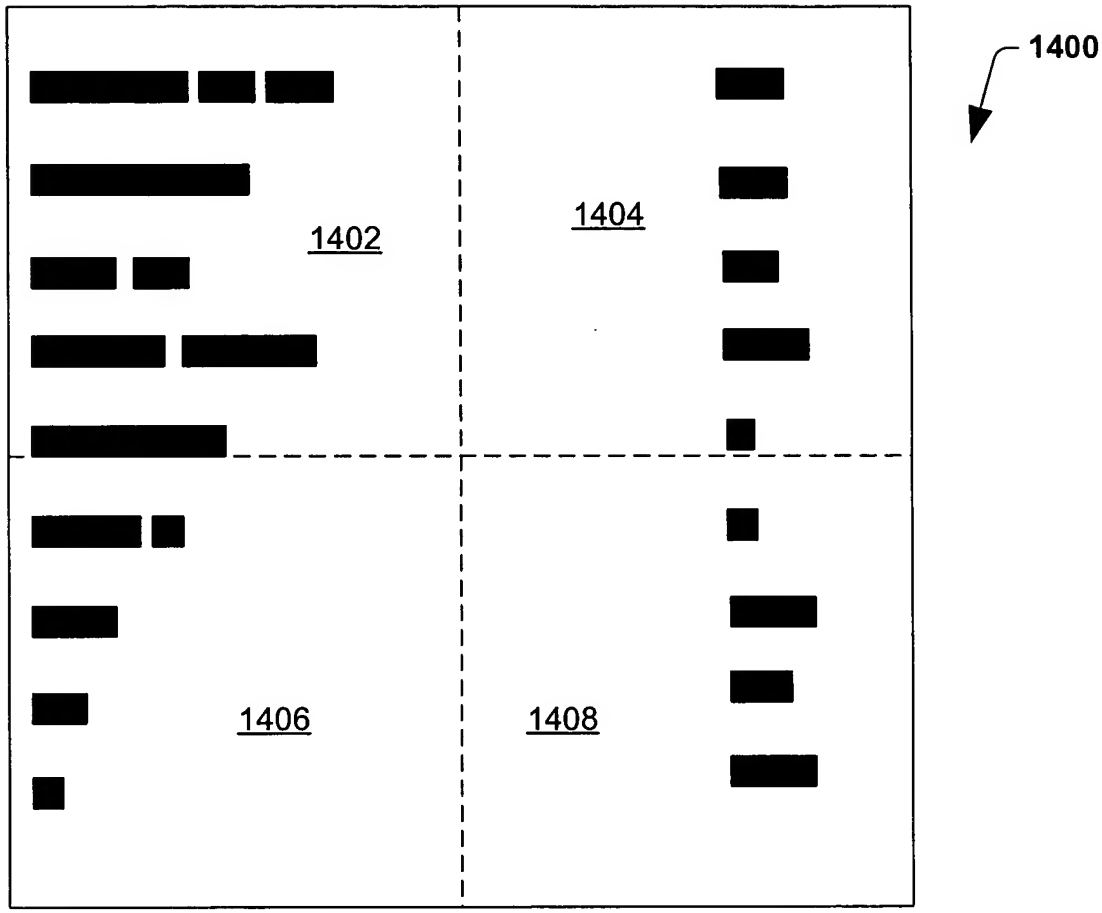

FIG. 14

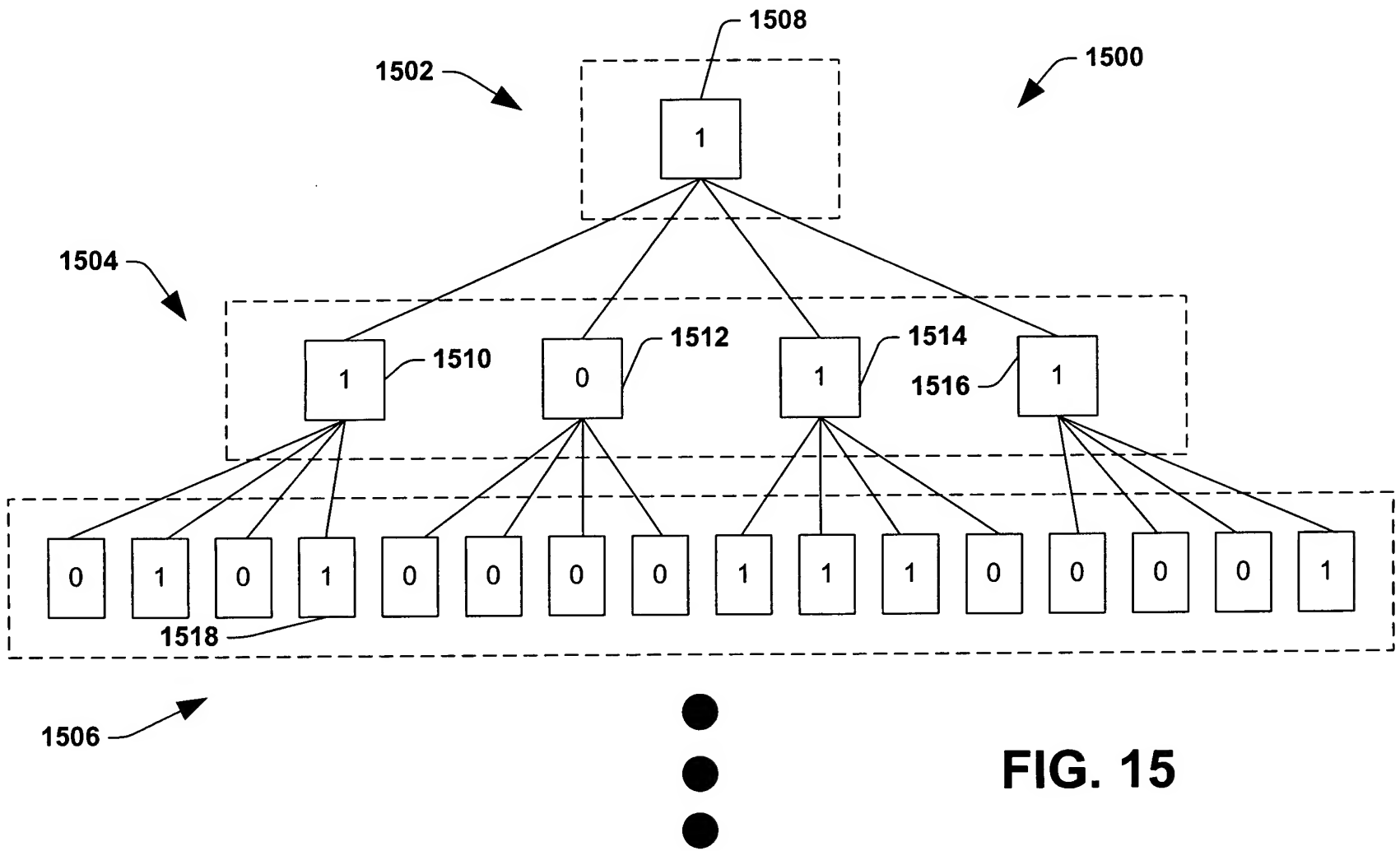
1508
1502
1500
1
1504
1
1510
0
1512
1
1514
1516
1
0
1
0
1
0
0
0
0
1
1
1
0
0
0
0
1
1518
1506

FIG. 15

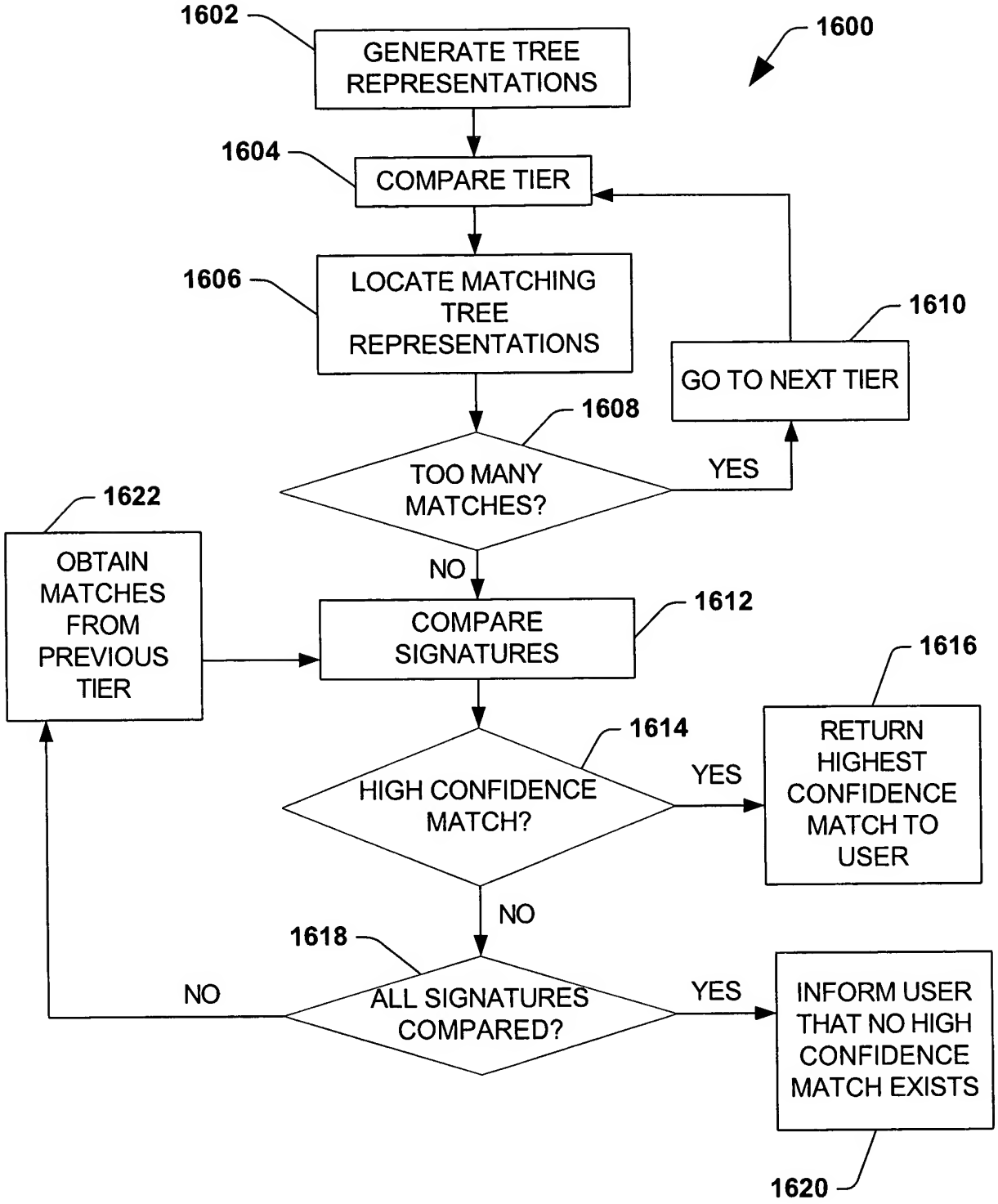
1602
GENERATE TREE REPRESENTATIONS
1600
1604
COMPARE TIER
1606
LOCATE MATCHING TREE REPRESENTATIONS
1610
GO TO NEXT TIER
1608
TOO MANY MATCHES?
YES
NO
1622
OBTAIN MATCHES FROM PREVIOUS TIER
1612
COMPARE SIGNATURES
1616
1614
HIGH CONFIDENCE MATCH?
YES
RETURN HIGHEST CONFIDENCE MATCH TO USER
NO
1618
NO
ALL SIGNATURES COMPARED?
YES
INFORM USER THAT NO HIGH CONFIDENCE MATCH EXISTS
1620

FIG. 16

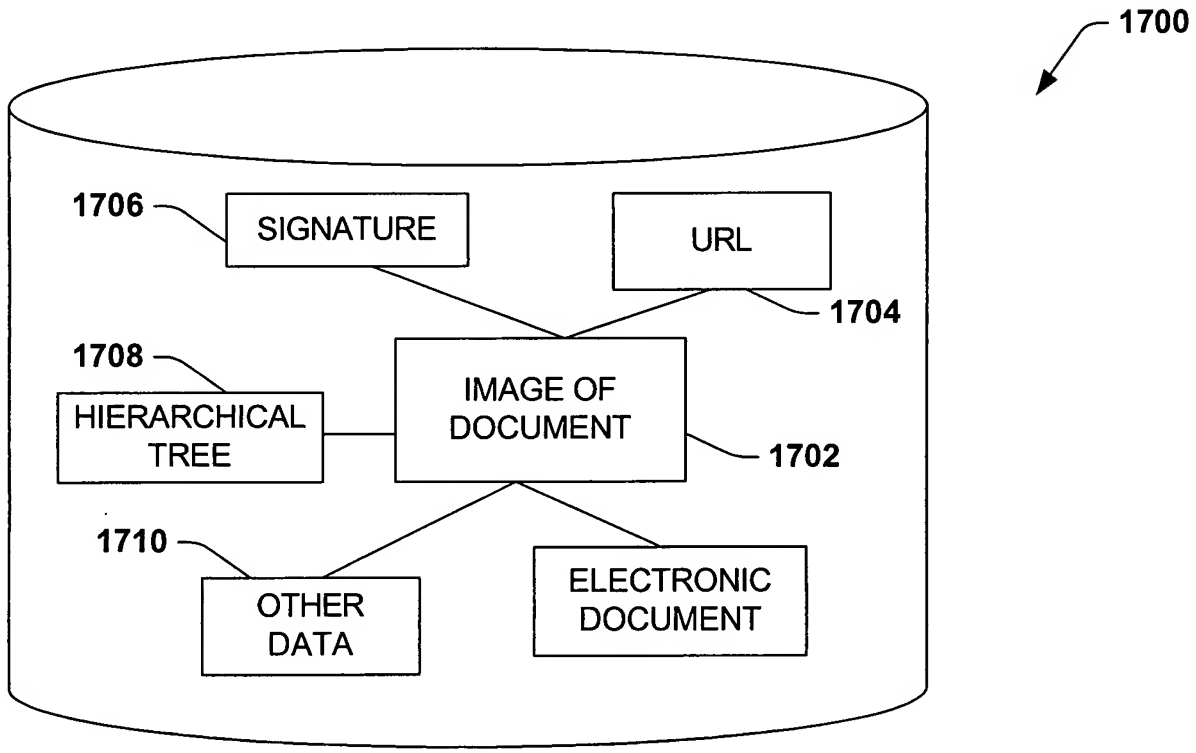
1700
1706
SIGNATURE
URL
1704
1708
HIERARCHICAL TREE
IMAGE OF DOCUMENT
1702
1710
OTHER DATA
ELECTRONIC DOCUMENT

FIG. 17

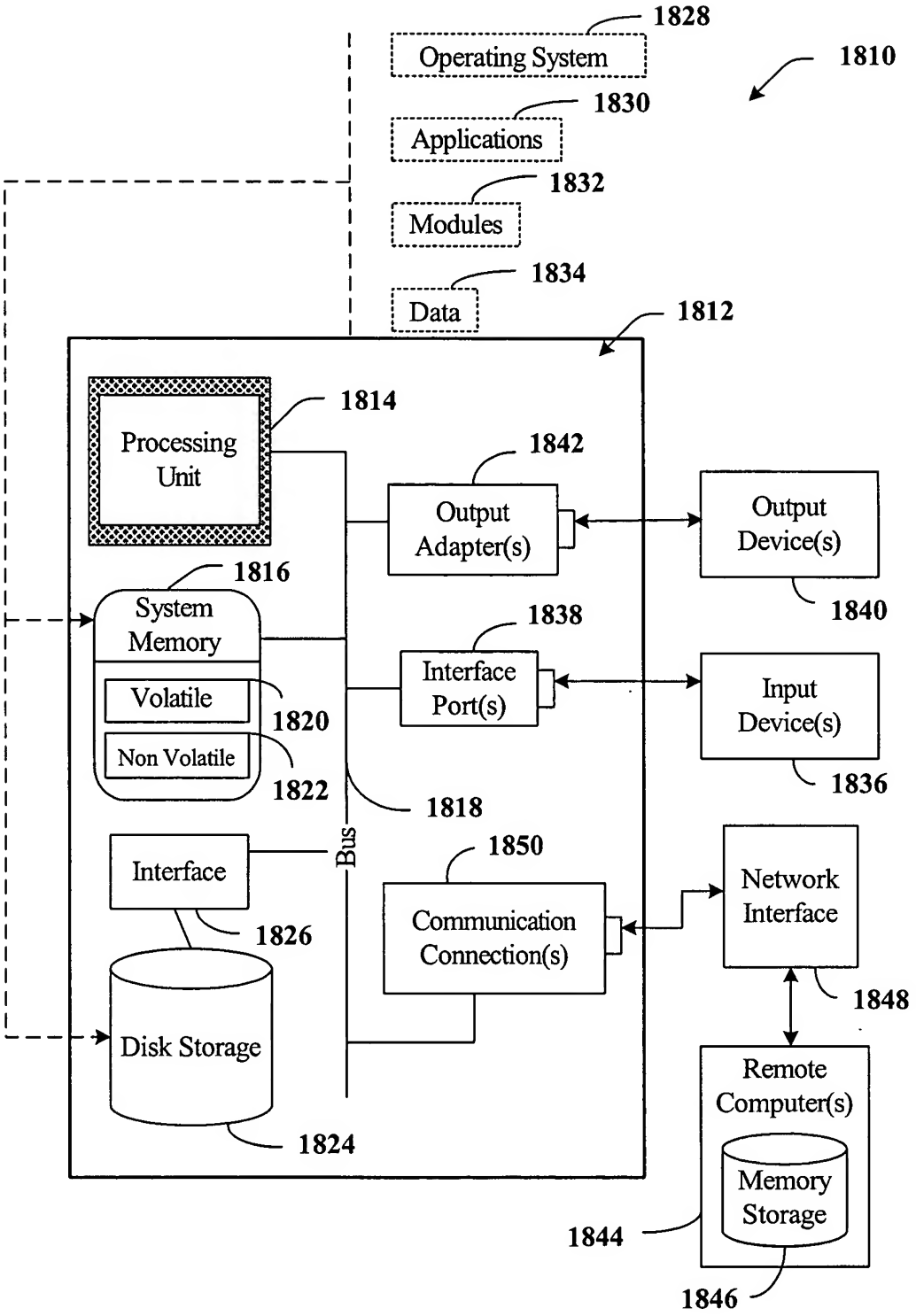
1828
Operating System
1810
1830
Applications
1832
Modules
1834
Data
1812
1814
Processing Unit
1842
Output Adapter(s)
Output Device(s)
1840
1816
System Memory
Volatile
1820
Non Volatile
1822
1838
Interface Port(s)
Input Device(s)
1836
1818
Bus
Interface
1826
1850
Communication Connection(s)
Network Interface
1848
Disk Storage
1824
Remote Computer(s)
Memory Storage
1844
1846

FIG. 18

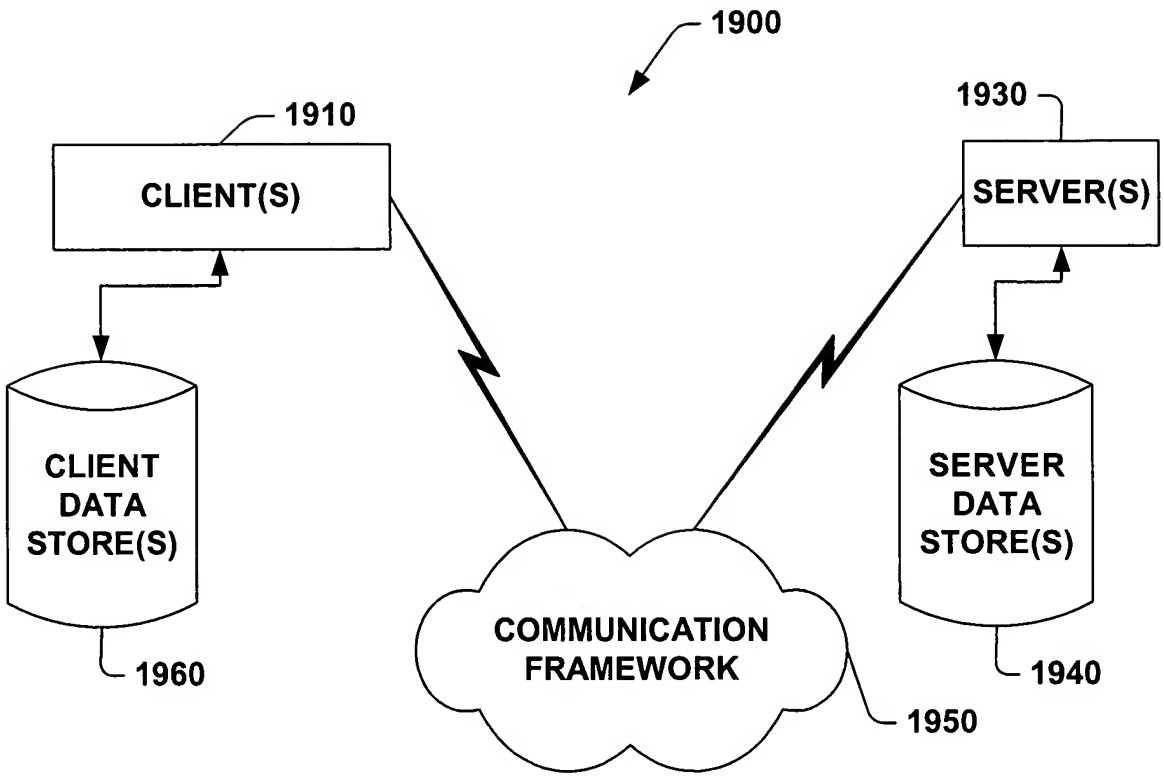
1900
1910
CLIENT(S)
1930
SERVER(S)
CLIENT DATA STORE(S)
1960
COMMUNICATION FRAMEWORK
1950
SERVER DATA STORE(S)
1940

FIG. 19